



# Topic Maps-based Semblogging with semblog-tm

Alexander Sigel, Lars Bünten & Marcus Frank University of Cologne, Germany Dept. of Information Systems & Information Management sigel@asigel.de | lars.b@gmx.de







Semantic blogging (semblogging) crosses blogs with the Semantic Web to enrich blog entries with more explicit and machineunderstandable metadata, relating both to structure and content.

This semantics is conveyed by semantic tags (semtags), connected by semantic associations (semrels).





Semblog-tm is a web-based application, realized as a plugin for the blojsom blogging engine, using Java and Tomcat with Velocity templates, with topic map operations via TMAPI, using TM4J and OKS as topic map engines, and with Axis for web services. A demonstrator is publicly available, and the source can be retrieved from sourceforge (semblog.sf.net).





### Outline

- 1. Motivation, Problem and Diagnosis
- 2. Prior art and related work
- 3. Approach
- 4. Demonstration of main functionalities as seen by sembloggers
- 5. System architecture
- 6. Discussion and Implications





Simple tagging is not enough Improve conventional blog tagging  $\rightarrow$  Folksonomies Lacks explicit structure (structural metadata) Lacks explicit semantics (content metadata) Tags not semantically grounded (only) names/strings) Tags not interrelated by typed semantic relations Difficult to syndicate different blog entries by subject

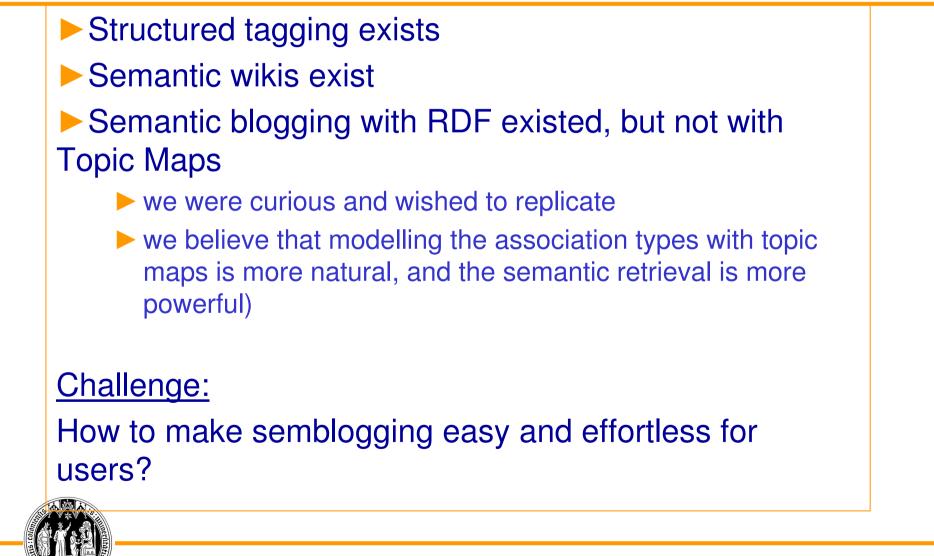




Conventional blog tagging → Folksonomies
 Tags not semantically grounded
 Tags not interrelated by typed semantic relations
 Enrich conventional blogging with more explicit, semantic, and machine-understandable metadata, relating both to structure and content









- 1. Seminal semblog system (by Steve Cayzer and colleages, HP/SWAD-E)
- 2. Haystack semblogging client (Karger & Quan)
- **3. Semblog** (Personal Knowledge Publishing Suite) (Ohmukai, Takeda, Numa and colleagues)
- 4. Semblogging research project (KMI: Sereno, Eisenstadt, Buckingham Shum)
- 5. Tagsocratic project (Avesani, Cova, Hayes, Massa)
- 6. semiBlog (semblogging to publish desktop data)





### Early ideas:

- Jack Park
- Dmitry Bogachev
- a playground implementation with OKS kindly provided by Lars Marius Garshol

### Recently:

 Rajbhandari, Andres, Naito & Wuwongse on semantic-augmented support in spatial-temporal multimedia blog management
 (@TMRA06, 2006-10-12, http://www.topincs.com/tmra/2006/?190)





Our analysis of the **six semblogging approaches and systems** identified reveals that:

▶ all are RDF-based

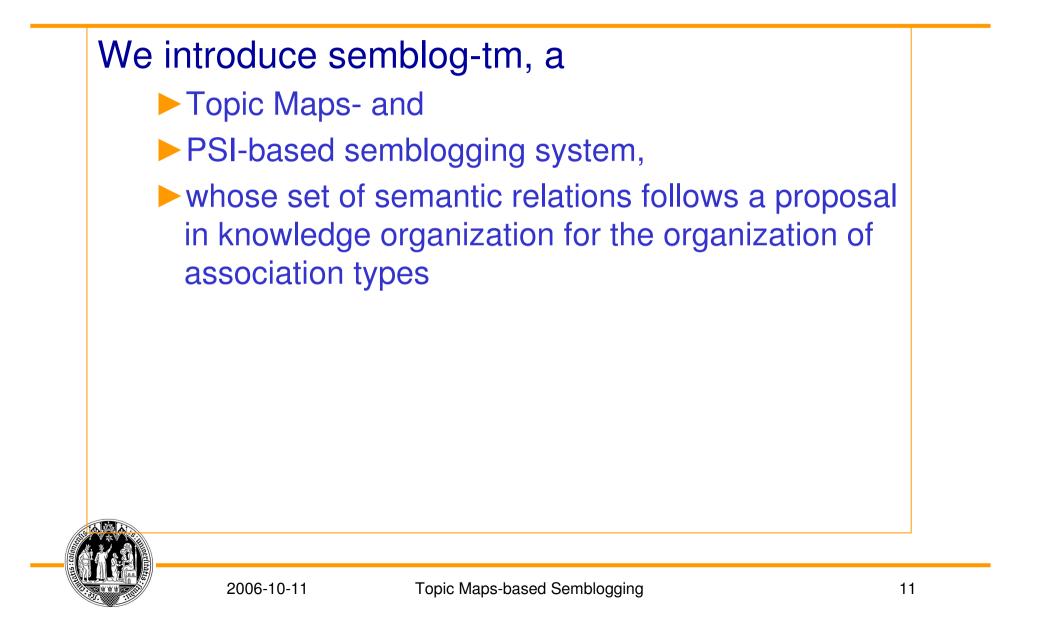
suffer from not using Published Subjects as proxies for subjects, and

do not employ semantic relations motivated by knowledge organization.

When we started in 2005, there was **no open effort on a topic maps-based semblogging system, and no research whatsoever** (see open space TMRA05 report)









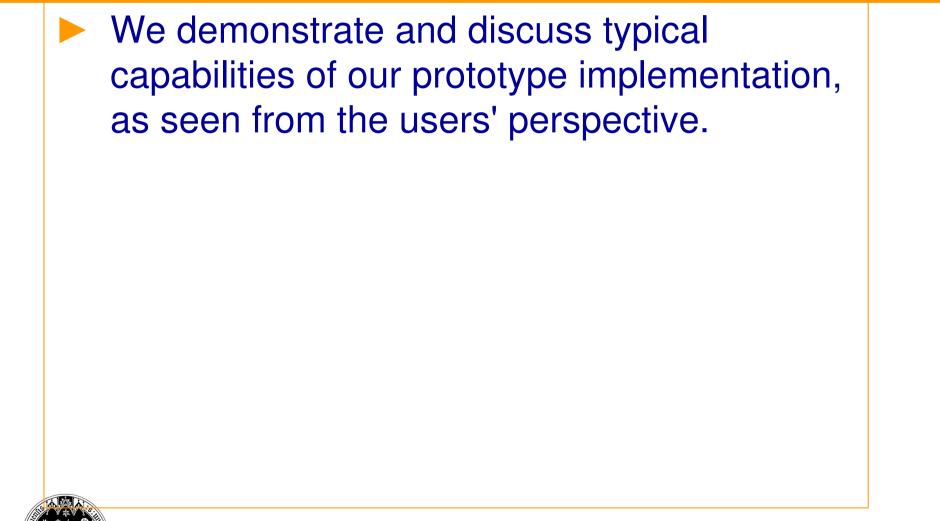
Starting from an application scenario with several sembloggers and servers, basic requirements are derived from the **three main use cases**:

(1) **Managing lightweight ontologies**: Defining ontologies (semtags connected by semantic associations).

(2) Semblogging: Connecting semblog entries with entries from those ontologies to attach semantics
(3) Managing aggregated semblogging data: Provide other semantically-enabled systems with topic map fragments.









#### #1 Nick blogging conventionally



Donnerstag, 28 September 2006 Attending TMRA06 (1)	« September »
After some problems with my flight, I finally arrived in Leipzig to attend TMRA06, and made way to Villa Ida. Nice new building. () In the spirit of this year's motto, Leveraging the Semantics, Steve Pepper proposed Public	Sun Mon Tue         Wed Thu         Fri         Sat           my         1         2         1         2           3         4         5         6         7         8         9           10         11         12         13         14         15         16           17         18         19         20         21         22         23
Resource Identifiers (PRIs), the successor of PSIs, to aid semantic interoperability. () In the demonstration session, a topic maps-based semblogging system called semblog was	24 25 26 27 28 29 30
shown. ()	About: Nick Name Links Administration Query own server Query all known servers Export XTM
This blog entry has the following semantics ( <u>edit</u> ): • <u>is written by semblogger_Nick</u> Formal metadata	Categories: / Syndication:
© Blog of Nick Name	Theme by Asual. Valid XHTML & CSS.

13. 38世世世 75



	Leveraging the Semantics 1	Topic Maps-based Portals and Information Retrieval	
	Chair: Roben, Barta <b>Room: Schille</b> r	Chair: Gabriel Hopmans Room: Everth + Peterhans	
9.00 - 10.30	Thea Miller, Hendrik Thomas	Prof. Sam Gyun Oh, Oknam Park	
	Indices, Meaning and Topic Maps: Some Observations 🖾 more	Design and Users' Evaluation of a Topic Maps-Based Korean Folk Music Retrieval System @ more	Discover a PSI on a webpage
	Lutz Maicher	Eszter Horvati	• •
	The Impact of Semantic Handshakes	Compass and more	(here: provided via Topincs)
	Patrick Durusau, Steven R. Newcomb	Henrik Laursen	
	The Essentials of the Topic Maps Reference Model (TMRM) @ more	Topic Maps Project at Royal Library, Copenhagen 🗹 more	
10.30 - 11.00	coffee	e break	
	Technical Issues of Topic Mapping	Leveraging the Semantics 2	
	Chair: Rani Pinchuk <i>Room: Schiller</i>	Chair: Markus Überall <b>Room: Everth + Peterhans</b>	
11.00 - 12.00	Graham Moore, Kal Ahmed, Andrew Brodie	Quintin H.J.F. Siebers, Peter-Paul Kruijsen	
	Topic Map Objects C more	Implementing Inference Rules in the Topic Maps Model 🕢 more	
	Robert Cerny	Lars Heuer	
	Topincs: Topic Maps, REST and JSON @	Towards converting the Internet into topic maps 🖉	



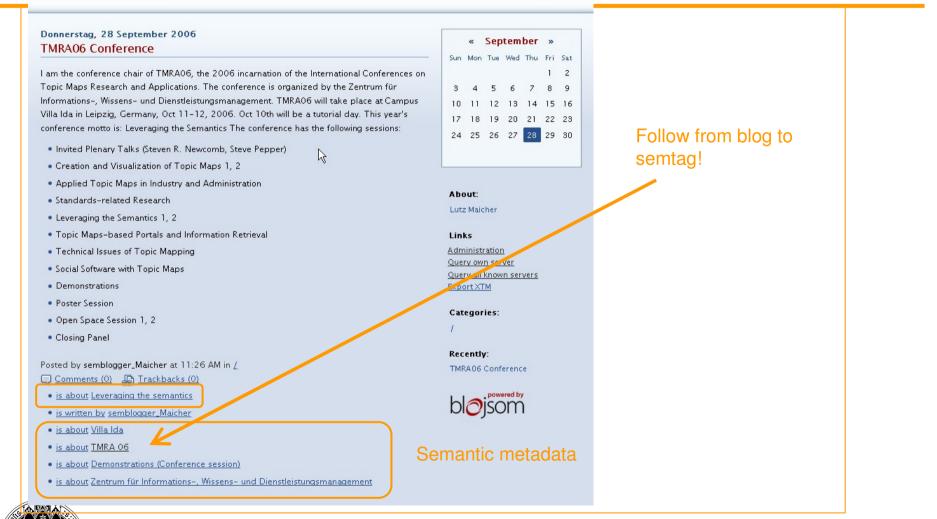
#3 Nick discovering other blog entries#4 Nick discovering semantics in other semblogger's blogs#5 Nick discovering semtags in Lutz Maicher's blog



#### Blog of Lutz Maicher This is my personal blog

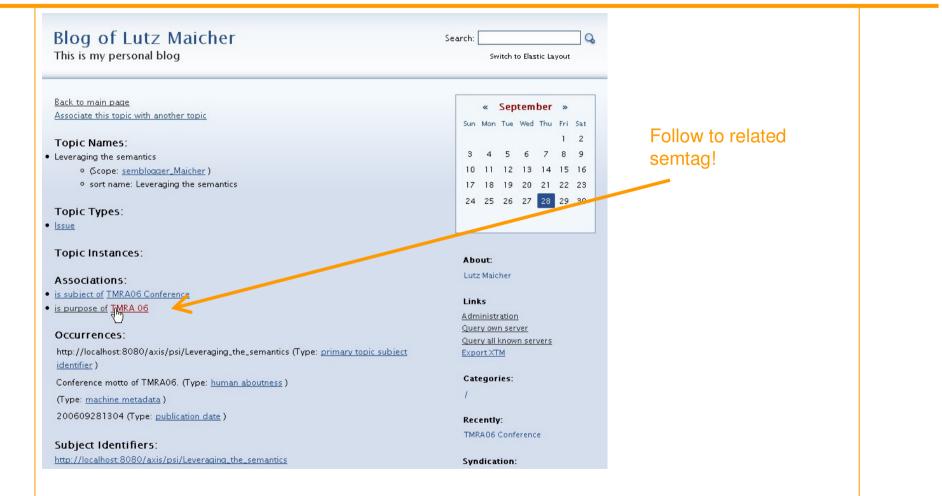


Q



Blog of Lutz Maicher This is my personal blog	Search: Switch to Elastic Layout		200
Back to main page Associate this topic with another topic	« September »	Ontology level:	
Associate this topic with another topic	Sun Mon Tue Wed Thu Fri Sat	<u>Ontology level.</u>	
Topic Names:	1 2		
• TMRA 06	3 4 5 6 7 8 9	Somtag	
<ul> <li>(Scope: <u>semblogger_Maicher</u>)</li> <li>sort name: TMRA 06</li> </ul>	10 11 12 13 14 15 16	Semtag	
o sort name. TMKA 00	17 18 19 20 21 22 23 24 25 26 27 <mark>28</mark> 29 30	(concept)	
Topic Types:	24 25 26 27 20 29 50		
• Issue			
Tanis Instances			
Topic Instances:	About:		
Associations:	Lutz Maicher		
is instance of Conference	Links		
<ul> <li>is preferential synonym used for International Conference on Topic Maps Rese</li> </ul>	arch and Application: Administration	Semtag-semtag	
<ul> <li>is instrumental for Leveraging the semantics</li> <li>descends from (is descendant of) TMRA05</li> </ul>	Query own server	relations	
v is subject of TwRA06 Conference	Query all known servers Export XTM	relations	
Occurrences:	Categories:		
200609281151 (Type: <u>publication date</u> )	T		
TMRA06 is the second conference in the TMRA conference series. TMRA – the			
conference series on Topic Maps Research and Applications - is a scientific a forum whose main object is connecting the key players in the Topic Maps con			
you will find researchers, users in government and industry, as well as the ver	innointy. Here		
luminaries, and the standards creators gathered for an exchange of ideas in	· · · · · ·		
setting. TMRA is where new challenges in Topic Maps are identified and oper tackled. (Type: <u>human aboutness</u> )	issues XML RSS ATOM		
(Type: <u>machine metadata</u> )	bla:		
http://www.tmra.de/psi/tmra06 (Type: <u>primary topic subject identifier</u> )	blojsom		
Subject Identifiers			
Subject Identifiers: http://www.tmra.de/psi/tmra06			













#### Three semtags: Leipzig Zentrum für Villa Ida Students of the University of Leipzig Students of the University of Leipzig Students of the University of Leipzig Syndicated blog by students of the University of Leipzig Syndicated blog by students of the University of Leipzig Syndicated blog by students of the University of Leipzig Back to main page Back to main page Back to main page Associate this topic with another topic Associate this topic with another topic Associate this topic with another topic Topic Names: Topic Names: Topic Names: • Villa Ida • Leipzig • Zentrum für Informations-, Wissens- und Dienstleistungsmanagement • (Scope: semblogger\_uleipzig) (Scope: semblogger\_uleipzig.) • (Scope: semblogger\_uleipzig ) o sort name: Villa Ida • sort name: Leipzig • sort name: Zentrum für Informations-, Wissens- und Dienstleistungsmanage Topic Types: Topic Types: Topic Types: • Issue Issue Issue Topic Instances: Topic Instances: Topic Instances: Associations: Associations: Associations: • is subject of What is Villa Ida? • is in place of (is geographic part of) Germany <u>A is cooperating with B</u> University Leipzig • is in place of (is geographic part of) Laipzig • is place for (is geographic whole for) University Leipzig • is subject of What is the Zentrum für Informations-, Wissens- und Dienstleistungsma • is place for (is geographic whole for) Villa Ida is in place of (is geographic part of) Leipzig Occurrences: • is place for (is geographic whole for) Zentrum für Informations-, Wissens- und Dienstleist, 200609281118 (Type: publication date ) 2 Occurrences: "Villa Ida" is a historical building in Leipzig. The Villa Ida is surrounded by the Occurrences: 200609281107 (Type: publication date ) Mediencampus Leipzig (Campus Villa Ida), an institutional part of the Leipzig Media 200609281046 (Type: publication date ) http://www.ziwd-leipzig.de (Type: primary topic subject identifier ) Foundation. (Type: human aboutness ) http://localhost:8080/axis/psi/Leipzig (Type: primary topic subject identifier ) Zentrum für Informations-, Wissens- und Dienstleistungsmanagement (Type: humar http://www.leipziger-medienstiftung.de/villa/ (Type: primary topic subject identifier ) The town Leipzig in Germany. (Type: human aboutness.) aboutness) (Type: machine metadata) (Type: machine metadata ) (Type: machine metadata ) Subject Identifiers: Subject Identifiers: Subject Identifiers:



#### #8 Nick following semtags in Steve Pepper's blog (1)

#### Steve Pepper's blog on Subject Identity

Advocating Subject Identity

Donnerstag, 28 September 2006

Switch to Elastic Layout

« September »

12 13 14 15 16

5 6 7 8 9

**Recently:** 

On PRIs and PSIs

Syndication:

XML RSS ATOM

bloisom

11

0

1 2

Search:

On PRIs and PSIs Sun Mon Tue Wed Thu Fri Sat A PRI (Public Resource Identifier) is an http: IRI that has been minted for the express purpose of identifying some subject. з PRIs are the successors of PSIs (Published Subjects). 10 17 18 19 20 21 22 23 Published Subjects are the successors of Public Subjects. 24 25 26 27 28 29 30 At TMRA06 I'll present "A Proposal for Public Resource Identifiers" Here is the abstract: About: The only way to achieve semantic interoperability is by making it possible to express what subjects a piece of information is about in such a way that it can be connected with other Steve Pepper pieces of information that are about the same subject. Because of the fragility of names, this requires the use of globally unique identifiers for subjects of discourse. This paper describes Links the requirements on a mechanism for defining and assigning unique global identifiers for Administration arbitrary subjects on the World Wide Web and presents the case for Public Resource Identifiers Query own server Query all known servers (PRIs) as the best solution to this problem. Export XTM Posted by semblogger\_Pepper at 1:20 PM in / Categories: Comments (0) 🔓 Trackbacks (0)

This blog entry has the following semantics (edit):

- is about Published Subject Indicator
- is about Published Subject
- is about TMRA 06
- is about Published Subject Identifier
- is about PSI
- is about Public Subject
- is about Public Resource Id ntifier
- is written by semblogger\_Pepper



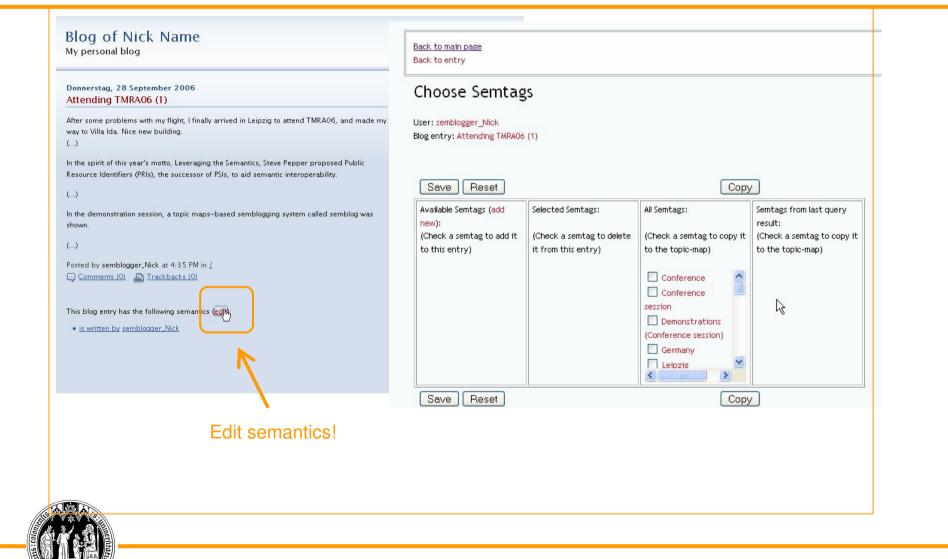


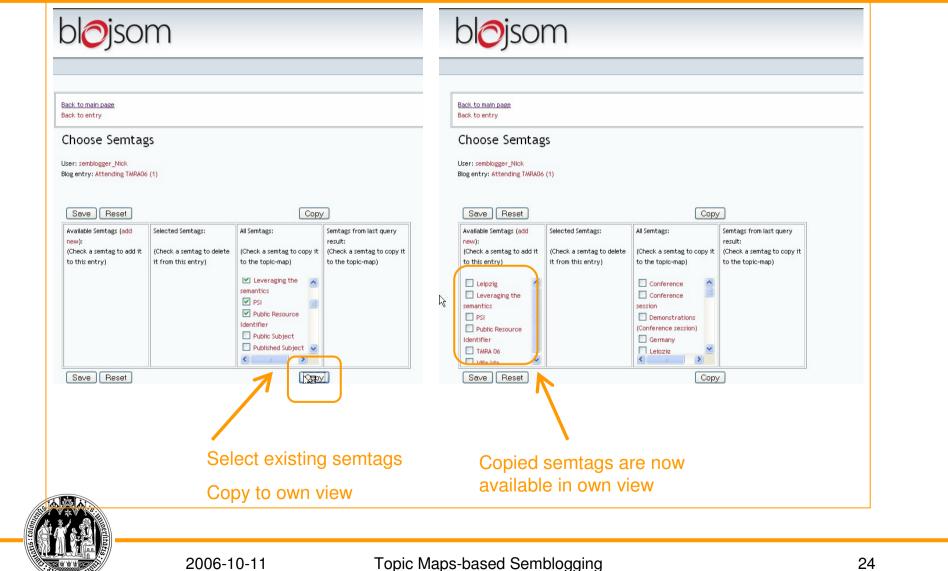


Public Resource Identifier	Published Subject	Public Subject
Steve Pepper's blog on Subject Identity Advocating Subject Identity	Steve Pepper's blog on Subject Identity Advocating Subject Identity	Steve Pepper's blog on Subject Identity Advocating Subject Identity
Back to main page Associate this topic with another topic	<u>Back to main page</u> Associate this topic with another topic	<u>Back to main page</u> Associate this topic with another topic
Topic Names: Public Resource Identifier • (\$cope: <u>semblogaer, Pepper</u> ) • sort name: Public Resource Identifier (PRI)	Topic Names: • Published Subject • (Scope: <u>semblogger_Pepper</u> ) • sort name: Published Subject	Topic Names: • Public Subject • (Scope: <u>semblogger_Pepper</u> ) • sort name: Public Subject
Topic Types: Issue	Topic Types: • Issue	Topic Types: • <u>Issue</u>
Topic Instances:	Topic Instances:	Topic Instances:
Associations: descends from (is descendant of) Published Subject is beneficial for Semantic Interoperability is subject of On PRIs and PSIs	Associations: • descends from (is descendant of) Public Subject • is composed of (has as matter) Published Subject Identifier • is predecessor of Public Resource Identifier • is subject of On PRIs and PSIs	Associations: • is predecessor of Published Subject • is subject of On PRIs and PSIs Occurrences:
Occurrences: http://localhost:8080/axis/psi/Public_Resource_Identifier (Type: <u>primary topic subject</u> <u>identifier</u> ) Concept for subject identification as proposed by Steve Pepper in his talk "A Proposal for	• <u>is composed of (has as matter)</u> Published Subject Indicator Occurrences: 200609281322 (Type: <u>publication date</u> )	200609281321 (Туре: <u>publication date</u> ) (Туре: <u>machine metadata</u> ) (Type: <u>human aboutness</u> )
Public Resource Identifiers" at TMRA06. (Type: <u>human aboutness</u> ) 200609281326 (Type: <u>publication date</u> )	http://localhost:8080/axis/psi/Published_Subject (Type: <u>primary topic subject identifier</u> ) Refinement of Public Subjects during XTM specification. (Type: <u>human aboutness</u> ) (Type: machine metadata )	http://localhost:8080/axis/psi/Public_Subject (Type: primary topic subject identifie Subject Identifiers: http://localhost:8080/axis/psi/Public_Subject

#### #10 Nick editing semtags





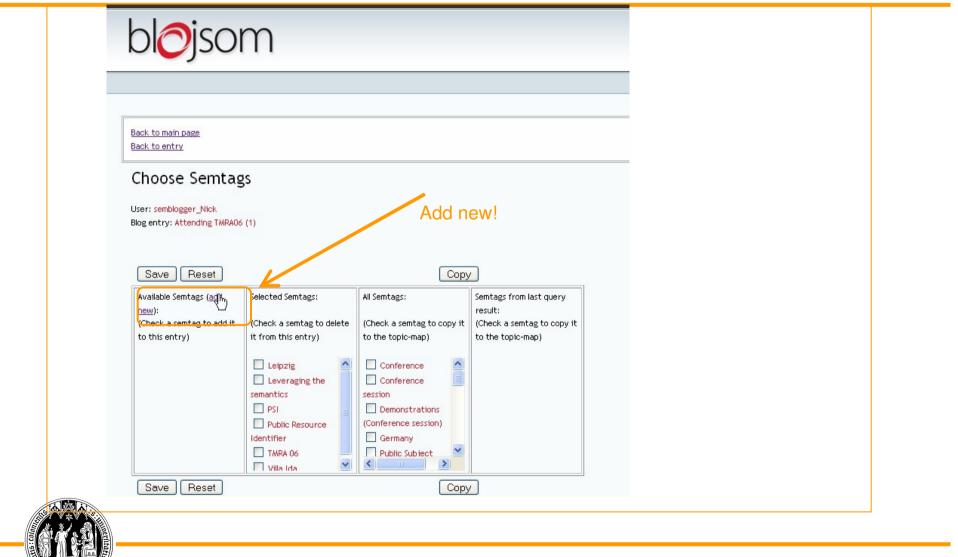


Back to main page		Back to main page
Back to entry		Back to entry
Choose Semtags		Choose Semtags
User: semblogger_Nick Blog entry: Attending TMRA06 (1)		User: semblogger_Nick Blog.entry: Attending TMRA06 (1)
Save Reset	Сору	Save Reset Copy
Available Semtags (add Selected	Semtags: All Semtags: Semtags from last query	Available Semtags (add Selected Semtags: All Semtags: Semtags from last query
	semtag to delete (Check a semtag to copy it (Check a semtag to copy it check a semtag to copy it is the testing and it is the testic and it is the testic and it is the testing	
to this entry) it from the lit	nis entry) to the topic-map) to the topic-map)	to this entry) it from this entry) to the topic-map) to the topic-map)
Leveraging the semantics	Conference	Leveraging the Conference
PSI     Public Resource	Demonstrations	PSI Demonstrations
Identifier	(Conference session)	Dublic Resource     (Conference session)       Identifier     Germany
TMRA 06     Villa Ida     V	C Leiozia	Villa Ida
Saye Reset	Сору	Sake Reset Copy
K		$\checkmark$
		1
	Select from available	Selected semtags now
	semtags in own view	attached to semblog entry

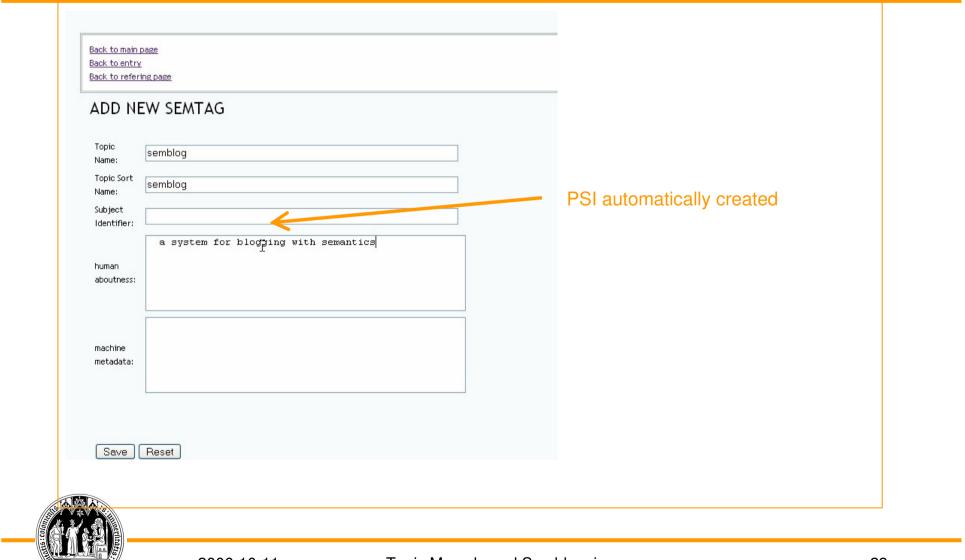
#### #12 Nick looking at the semtags in his entry



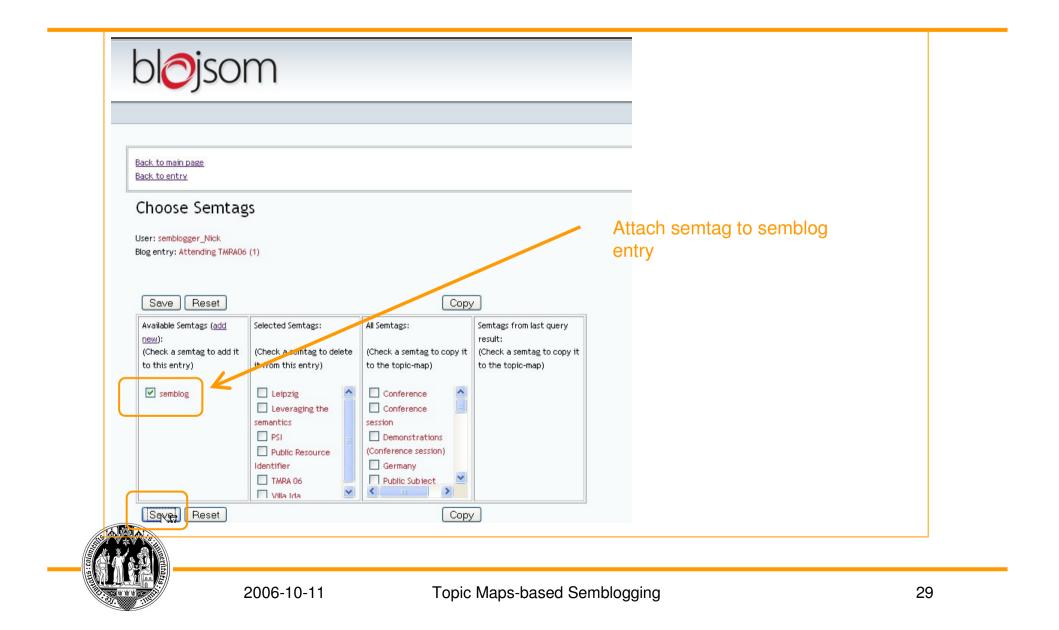
'MRΔ

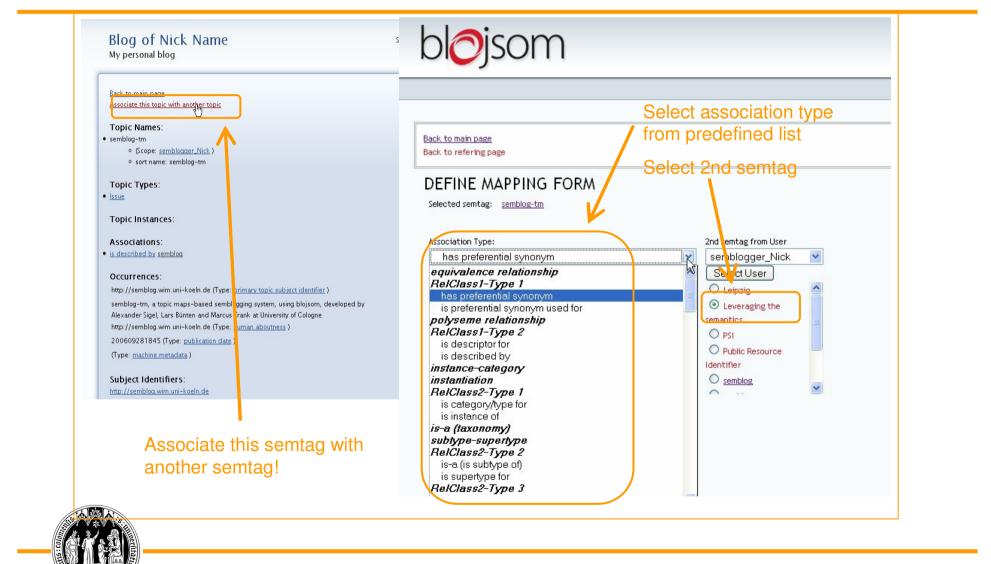


#### #14 Nick defining his own semtag (2)



TMRA







<u>Back to main page</u> Back to refering page			
DEFINE MAPPING FORM Selected semtag: <u>semblog-tm</u>		The chosen mapping is:	
Association Type: is beneficial for	2nd semtag from User semblogger_Nick Select User Leipzig Leveraging the semantics PSI Public Resource Identifier Semblog	Semblog-tm is beneficial for Leveraging the Semantics	

#### #21 Narrator explaining semantic retrieval



Available semtags for all users [topicmap_0, to semblogger_uleipzig, user1]:	picinap_2, iais, semblogger_mick, (	ueraun,	3	4	5	6	7	8	
O Conference		~	10	11	12	13	14	15	;
O Conference session			17	18	19	20	21	22	:
O Demonstrations (Conference session)			24	25	26	27	28	29	,
O Germany									
			_						
O Leveraging the semantics									
OPSI			Abo	out:					
O Public Resource Identifier			Nick	Nam	ne				
O Public Subject									
O Published Subject			Lini						
O Published Subject Identifier			197222000		ratior n ser	12.			
O Published Subject Indicator	6		2.565.55			n ser	vers		
O <u>Semantic Interoperability</u>	10		18-11-11-11-11-11-11-11-11-11-11-11-11-1				13		
O <u>semblog</u>									
O <u>semblog</u>			Rec	enth	V:				
O <u>semblog-tm</u>				ndin	g TMP	RA06			
Otmra			(1)						
⊙ <u>TMRA 06</u>			5ve	dica	tion:				
O TMRA05		~	зул XML						

Query Reset Complete results only servers in list Query results for TMRA 06 : 1 Attending TMRA06 (1) 2 On PRIs and PSIs 3 TMRA06 Conference





#### #22 Nick querying semtags by all sembloggers on the **same** semblog server

Back to main page	« September »	
	Sun Mon Tue Wed Thu Fri Sat	
Available semtags for all users [lars, semblogger_Nick, default, semblogger_uleipzig,	1 2	
user1]:	3 4 5 6 7 8 9	
Germany	Blog of Nick Name	Search:
	My personal blog	Switch to Elastic Layout
Leveraging the semantics	, p	
Public Resource Identifier	Back to main page	
semblog		« September »
semblog		Sun Mon Tue Wed Thu Fri Sat
semblog-tm	Available semtags for all users [lars, semblogger_Nick, default, semblogger_uleipzig,	1 2
	user]]:	3 4 5 6 7 8 9
TMRA 06	Germany	10 11 12 13 14 15 16
University Leipzia		17 18 19 20 21 22 23
□ <u>Villa Ida</u>	Leveraging the semantics	24 25 26 27 28 29 30
Zentrum für Informations-, Wissens- und Dienstleistungsmanagement		
	Public Resource Identifier	
	semblog	
Query Reset	semblog	About:
	semblog-tm	Nick Name
Query	TMRA 06	Links
a,	University Leipzia	Administration Query own server
	Uilla Ida	Query all known servers
	Zentrum für Informations-, Wissens- und Dienstleistungsmanagement	Export XTM
		Categories:
		(
	Query Reset Ouery results for TMRA 06 : Follow result	·
		Recently:
	1 Attending TMEA06 (1)	Attending TMRA06
		(1)

TMRA

#### #24 Nick querying semtags by **all sembloggers on all known semblog servers** for the semtag TMRA06 (display complete results)

semblogger_uleipzig, user1]:	10 11 12 13 14 15 16
O <u>Conference</u>	17 18 19 20 21 22 23
O <u>Conference session</u>	24 25 26 27 28 29 30
O Demonstrations (Conference session)	24 25 26 27 29 30
O <u>Germany</u>	
O <u>Leipziq</u>	
O <u>Leveraging the semantics</u>	
O <u>PSI</u>	About:
O <u>Public Resource Identifier</u>	Nick Name
O <u>Public Subject</u>	Links
O <u>Published Subject</u>	Administration
O Published Subject Identifier	Query own server
O Published Subject Indicator	Query all known servers
O <u>Semantic Interoperability</u>	Export result to XTM
O <u>semblog</u>	
O <u>semblog</u>	Recently:
O <u>semblog-tm</u>	Attending TMRA06
O <u>tmra</u>	(1)
⊙ <u>TMRA 06</u>	Syndication:
O <u>tmraos</u>	XML RSS ATOM
	powered by
	DIOJSOM
Query Reset 🗹 complete results 🗖 only servers in list	
Query results for <u>TMRA 06</u> :	
1 <u>Attending TMRA06 (1)</u>	
2 <u>On PRIs and PSIs</u>	
<sup>3</sup> <u>TMRA06 Conference</u> Follow result	



TMRA

### #25 Nick displaying the results and exporting them to .xtm



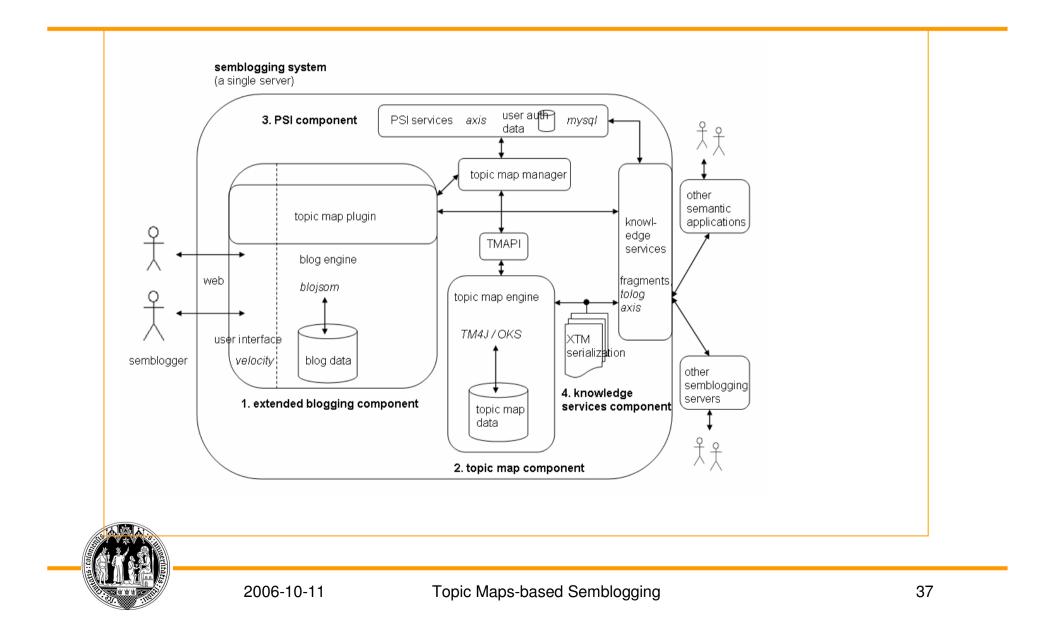
Topic Types:	
Blog entry	Mit dieser XML-Datei sind anscheinend keine Style-Informationen verknüpft. Nachfolgend wird die Bar
Topic Instances:	
	- <topicmap xml:base="http://localhost:8080/blojsom_2/blog/semblogger_Nick/usermap.xtm"></topicmap>
Associations:	- <topic id="id178"></topic>
• is about PSI	- <instanceof></instanceof>
is about Public Subject	<topicref xlink:href="#id2"></topicref>
is about Published Subject Indicator	
is about Published Subject Identifier	- <subjectidentity></subjectidentity>
is about Published Subject	<subjectindicatorref xlink:href="http://www.schmitz-esser.de/ICLO/v0.01/rt_enhancer"></subjectindicatorref>
is about Public Resource Identifier	
is written by semblogger_Pepper	-   
	  baseNameString>enhancer
• <u>is about TMRA 06</u> The result	
Occurrences:	- <topic id="id75"></topic>
Thu Sep 28 13:20:51 CEST 2006 (Type: Blog entry publication date )	- <instanceof></instanceof>
•	<topicref xlink:href="#id2"></topicref>
http://localhost:8080/blojsom/blog/semblogger_Pepper/?permalink=On-PRIs-and-PSIs.htr	
(Type: <u>Permalink</u> )	- <subjectidentity></subjectidentity>
	<subjectindicatorref xlink:href="http://www.schmitz-esser.de/ICLO/v0.01/rt_regulated"></subjectindicatorref>
A PRI (Public Ressource Identifier) is an http: IRI that has been minted for the express	
purpose of identifying some subject.	- <basename></basename>
PRIs are the successors of PSIs (Published Subjects).	   
Published Subjects are the successors of Public Subjects.	The VTM expect
At TMRA06 I'll present "A Proposal for Public Resource Identifiers"	<pre> The XTM export</pre>
Here is the abstract:	- <subjectidentity></subjectidentity>
	<subjectindicatorref xlink:href="http://psi.wim.uni-koeln.de/blog-entry-content"></subjectindicatorref>
The only way to achieve semantic interoperability is by making it possible to express what	subjectIdentity>
subjects a piece of information is about in such a way that it can be connected with other	- <basename></basename>
pieces of information that are about the same subject. Because of the fragility of names, this	
requires the use of globally unique identifiers for subjects of discourse. This paper	
describes the requirements on a mechanism for defining and assigning unique global	



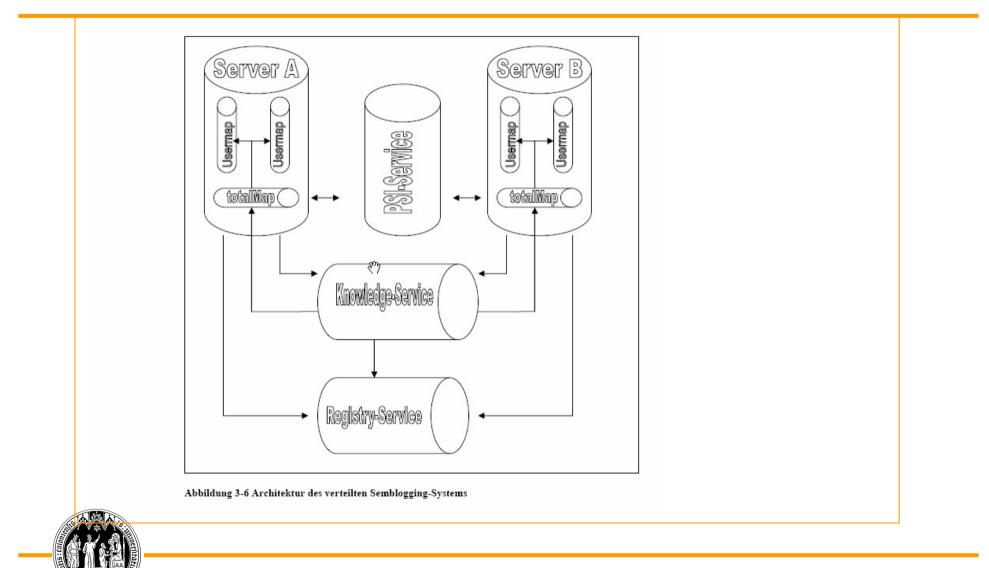




#### #26 Narrator explaining the general system architecture

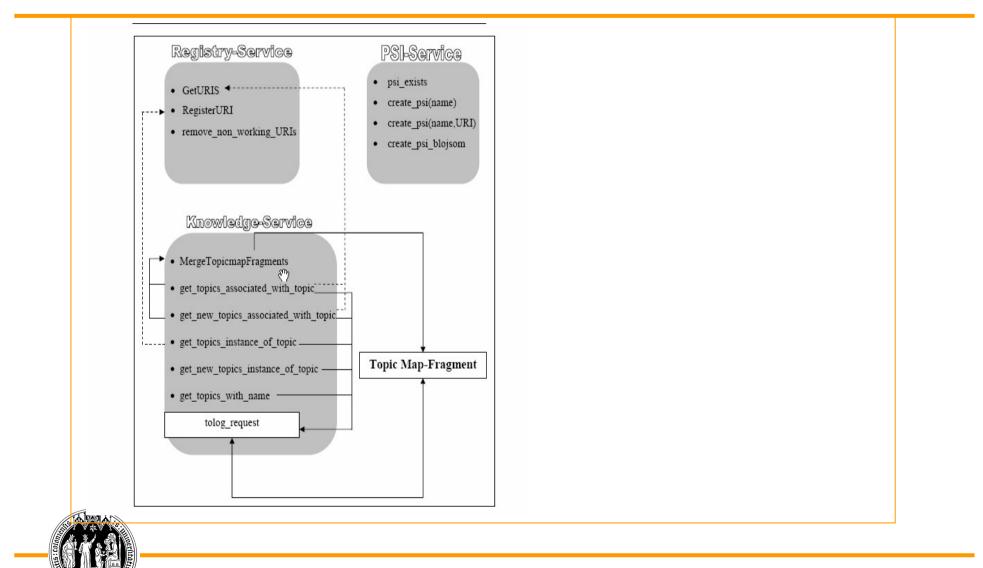


## #27 Narrator explaining scalability and distribution/P2P-like aspects



TMRA

#### #28 Narrator explaining semantic knowledge services





- Currently only a university research prototype, not in practical use, needs to move to being actively used by a community
- Needs some clean-up (e.g. GUI, code, naming restrictions, use of published subjects), and optimization
- Just one particular set of predefined association types (clearly needs refinement, e.g. rhetorical ones)
- No possibility to define application-specific subtypes of association types
- By design, only semtags can be copied, not semrels (associations)





- Instead of the association type Synonymy basename/scope would be more topic mappish
- PSI service could be more sophisticated, in the long run to be replaced by a third-party solution
- Use of TMAPI limits (e.g. no TMRAP, because TMRAP or tolog via TMAPI is not yet implemented)
- Needs OKS for tolog, because TM4J's tolog is not sufficient





- First open-source topic maps-based semblogging system
- The only semblogging system using PSIs
- The only semblogging system based on association types motivated and governed by Knowledge Organization theory
- First semblogging system with open semantic knowledge services, aggregating the data and providing them back to any topic map fragment consumers
   → cooperation with other PSI services!





- Expand actual use in semblogging communities
- Explore interoperability with other PSI and topic map aggregation services, with knowledge service consumers
- Explore P2P-like distribution further: e.g. TMRAP and semantic overlay in P2P networks with topic map data
- Further work under way on RDF/Sparql with topic maps/tolog interoperability in Semantic Knowledge Services, consuming semblog-tm services





